AD________________

Award Number: W81XWH-13-2-0010

TITLE: Semantic Web Technology for Mapping and Applying Clinical Functional Assessment Information

PRINCIPAL INVESTIGATOR:
Mark A. Musen, MD PhD

CONTRACTING ORGANIZATION:
Stanford University
Stanford, CA 94305-5479

REPORT DATE:
March 2014

TYPE OF REPORT:
Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE March 2014	**2. REPORT TYPE** Annual Report	**3. DATES COVERED** 15 February 2013 – 14 February 2014

4. TITLE AND SUBTITLE
Semantic Web Technology for Mapping and Applying Clinical Functional Assessment Information

5a. CONTRACT NUMBER
W81XWH-13-2-0010

5b. GRANT NUMBER
W81XWH-13-2-0010

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Mark A. Musen MD PhD

E/Mail: musen@stanford.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Medical School Office Building, Room X-215
1265 Welch Road
Stanford, CA 94305-5479

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We report on the progress of the project during the Feb 2013 – Feb 2014 period of the contract. During the year, we interviewed Madigan Army Medical Center collaborators to learn about their Integrated Disability Evaluation System. We found that the documentation related to functional assessment is in PDF and it's difficult to de-identify and export in structured format. We analyzed retention standards, Disability Benefit Questionnaires, the Military Occupation Specialties manual, and the Department of Veteran Affairs schedule for rating disabilities, and developed a framework for structuring and using functional assessment information. Within this framework, we modeled as OWL ontologies descriptions of functional assessments and their value sets and created data models for patient-specific data and for assessment forms such as DBQs. We map the functional assessments descriptors to ICF categories and qualifiers, and used the SWRL rules to transform CFA-CL coded data to ICF formats, when such transformations are possible.

15. SUBJECT TERMS
Functional Status; International Classification of Functioning, Disability, and Health; Disability Benefit Questionnaire; Integrated Disability Evaluation System

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU		**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Page

Introduction

In early 2013, the Stanford Center for Biomedical Informatics Research (BMIR) received a contract from TATRC to develop a "Common Language" for clinical functional assessment (CFA-CL). It is a two-year contract starting in February 2013 and terminating in February 2015. This document describes the status of the project at the end of the first year. It describes our findings on the DoD/VA's Integrated Disability Evaluation System (IDES), our revised project goals, our modeling approach, and the current problems. In the appendices, we summarize background information on the ICF and our prior work with the Social Security Administration (SSA) that is relevant to the current project.

Keywords

Functional Status; International Classification of Functioning, Disability, and Health; Disability Benefit Questionnaire; Integrated Disability Evaluation System

Overall Project Summary

We have named the project "FACSIMILE."1 Its initial goals, as described in our original proposal, include:

- Development of a "Common Language" (CFA-CL) for clinical function assessments that is grounded in International Classification of Functioning, Disability, and Health (ICF)
- Demonstration that data used in DoD/VA's Integrated Disability Evaluation System (IDES) can be coded in this common language
- Demonstration of uses of coded clinical function assessment data in the IDES process
- Creation of a prototype CFA semantic model in which categories of impairment are defined by constraint expressions consisting of the CFA-CL and ICF code stems, qualifiers, and qualifier values.

Our focus is IDES, through which DoD and VA providers and coordinators both evaluate a service member for fitness for service and determine a possible disability rating in parallel, thus reducing the required processing time for a disabled service member to begin receiving benefits. In the following, we describe the results and progress of the project in terms of the tasks outlined in the Statement of Work for the first year.

1. Analyze the functional requirements of tasks in the Integrated Disability Evaluation System (IDES) workflow where clinical functions are assessed, documented, stored, transmitted, and used.

In the first part of this project, we engaged in extensive consultation with colleagues in Madigan Army Medical Center to determine (1) the nature of clinical assessment information generated and used in IDES, and (2) opportunities to use coded clinical functional assessment information to inform decision-making in IDES.

In the IDES process, when the illness or injury of a service member fits the criteria defined in the medical fitness standards for retention and separation (e.g., Army Regulation 40-501[2]), and further treatment will not cause the member to meet medical retention standards or render them capable of performing the duties required by their office, grade, rank, and rating, the heath-care provider refers the service member to a Medical Evaluation Board (MEB) for the initiation of the IDES process and a Physical Evaluation Board Liaison Officer (PEBLO) is appointed for the service member. The PEBLO prepares and submits the case file to a VA Military Service Coordinator (MSC), who initiates VA processing of the case, schedules a medical exam, and sends the exam results to the PEBLO. The MEB providers use all available information to produce a narrative summary. The narrative summary, together with the service member's medical and service profiles and the history and treatment of the injury or illness, is used by the MEB to determine whether the member has a medical condition that is incompatible with continued military service in his or her current capacity. After a review process, the case file is sent to the Physical Evaluation Board (PEB) for determining the service member's fitness for service. An informal PEB makes the initial determination, and if the service member is found unfit, submits a

---

1 Functional-Assessment Coding for Semantic Interpretation of Military Impairment-Level Evaluation

2 The DBQs are VBA-21-0960M-14-ARE-Back.pdf, VBA-21-0960M-9-ARE-KneeLowerLeg.pdf, VBA-21-

request for disability ratings of all claimed conditions. A Disability Rating Activity Site issues a rating based on the findings of the VA medical examination. The process ends when all reviews and appeals have been processed and the disposition of the case is approved by the Physical Disability Agency (PDA) for the service member's return to duty or for the issuance of a VA's benefits decision letter.

We found that Madigan AMC's IDES data processing relies on PDF documents. The documents include narrative summaries in free text (Figure 1) or form-based documents that are accessible as PDFs (Figure 2). Because there is no coding scheme for clinical functional assessments (something that this project is addressing), functional assessment information in Madigan's IDES documentation is scattered in various narrative documents. ICD codes are the only structured data that are readily available. We procured an example of the dossier that is generated for a service member. It consists of 45 pages of mostly narrative notes that require significant time to redact and de-identify. The dossier provides many examples of clinical functional assessments (e.g., see highlighted text in Figure 1). However, it would take a herculean effort to convert such free text into coded data post hoc. Within the workflow of IDES as carried out at Madigan, we see no prospect of such structured coding being done. We concluded that it's unrealistic to expect that we can obtain a large sample of de-identified data from Madigan.

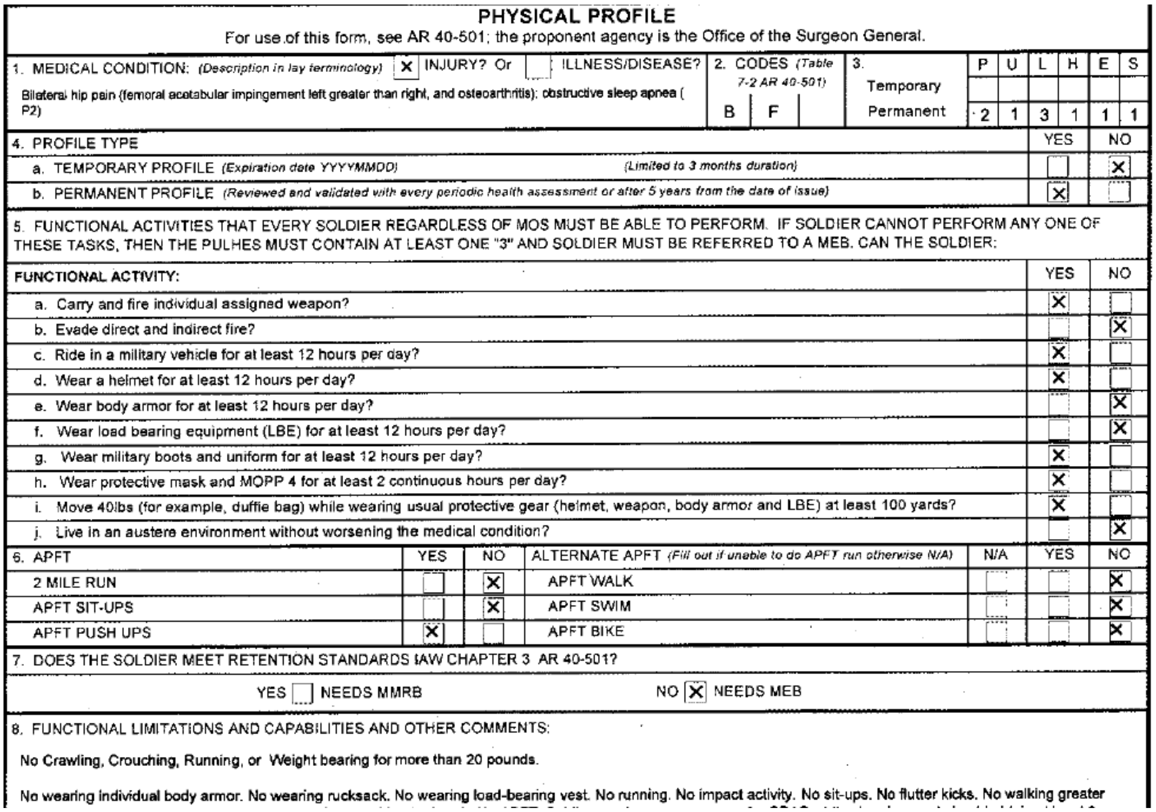
Furthermore, it's not clear what would be a good use case for the structured functional assessment information in Madigan's current IDES process. Madigan MEB physicians and a PEB officer emphasized to us in interviews that the retention decision is based on a holistic evaluation of many sources of information, including the service member's motivation and his/her superiors' assessments, rather than on any kind of structured assessment data. We struggled to find decision points where structured data can play a role in Madigan's IDES process.

Nevertheless, IDES is a complex and evolving process where a number of DoD and VA information systems interact with each other. We do not rule out the possibility of structured functional assessment data becoming useful in Madigan's IDES process in the future. We will be happy to revisit Madigan and apply the outputs of this project.

SPECIFIC HISTORY FOR: CHRONIC BILATERAL HIP CONDITION SECONDARY TO OSTEOARTHRITIS

The claimant reports being diagnosed with CHRONIC BILATERAL HIP CONDITION SECONDARY TO OSTEOARTHRITIS'. The condition has existed since 2005. The condition started after a 12 mile road march He reports the following symptom(s): pain. He indicates he does not experience weakness, stiffness, swelling, heat, redness, giving way, lack of endurance, locking, fatigability, deformity, tenderness, drainage, effusion, subluxation and dislocation. The claimant reports experiencing the following flare-ups as often as 7 time(s) per week and each time lasts for 1 day(s). From 1 to 10 (10 being the worst) the severity level is at 10. The flare-ups are occurring spontaneously. It is alleviated by rest, spontaneously and by Toradol. During the flare-ups he experiences the following functional impairment which is described as not able to run or walk fast and limitation of motion of the joint which is described as limited range of motion. He reports difficulty with standing/walking. The claimant describes can't stand or sit still for more than 3o minutes The treatment is Toradol, physical training, Tylenol

Figure 1. Example of functional assessment information embedded in narrative text of notes.



PHYSICAL PROFILE

For use of this form, see AR 40-501; the proponent agency is the Office of the Surgeon General.

1. MEDICAL CONDITION: *(Description in lay terminology)* [X] INJURY? Or [ ] ILLNESS/DISEASE?
Bilateral hip pain (femoral acetabular impingement left greater than right, and osteoarthritis); obstructive sleep apnea (P2)

2. CODES *(Table 7-2 AR 40-501)*: B | F

3.	P	U	L	H	E	S
Temporary						
Permanent	2	1	3	1	1	1

4. PROFILE TYPE	YES	NO
a. TEMPORARY PROFILE *(Expiration date YYYYMMDD)* *(Limited to 3 months duration)*		X
b. PERMANENT PROFILE *(Reviewed and validated with every periodic health assessment or after 5 years from the date of issue)*	X	

5. FUNCTIONAL ACTIVITIES THAT EVERY SOLDIER REGARDLESS OF MOS MUST BE ABLE TO PERFORM. IF SOLDIER CANNOT PERFORM ANY ONE OF THESE TASKS, THEN THE PULHES MUST CONTAIN AT LEAST ONE "3" AND SOLDIER MUST BE REFERRED TO A MEB. CAN THE SOLDIER:

FUNCTIONAL ACTIVITY:	YES	NO
a. Carry and fire individual assigned weapon?	X	
b. Evade direct and indirect fire?		X
c. Ride in a military vehicle for at least 12 hours per day?	X	
d. Wear a helmet for at least 12 hours per day?	X	
e. Wear body armor for at least 12 hours per day?		X
f. Wear load bearing equipment (LBE) for at least 12 hours per day?		X
g. Wear military boots and uniform for at least 12 hours per day?	X	
h. Wear protective mask and MOPP 4 for at least 2 continuous hours per day?	X	
i. Move 40lbs (for example, duffle bag) while wearing usual protective gear (helmet, weapon, body armor and LBE) at least 100 yards?	X	
j. Live in an austere environment without worsening the medical condition?		X

6. APFT	YES	NO	ALTERNATE APFT *(Fill out if unable to do APFT run otherwise N/A)*	N/A	YES	NO
2 MILE RUN		X	APFT WALK			X
APFT SIT-UPS		X	APFT SWIM			X
APFT PUSH UPS	X		APFT BIKE			X

7. DOES THE SOLDIER MEET RETENTION STANDARDS IAW CHAPTER 3 AR 40-501?
YES [ ] NEEDS MMRB NO [X] NEEDS MEB

8. FUNCTIONAL LIMITATIONS AND CAPABILITIES AND OTHER COMMENTS:
No Crawling, Crouching, Running, or Weight bearing for more than 20 pounds.
No wearing individual body armor. No wearing rucksack. No wearing load-bearing vest. No running. No impact activity. No sit-ups. No flutter kicks. No walking greater

Figure 2. Example of formed based information collection.

Locally, we interviewed Dr. Michael Tierney, a physician at the VA Palo Alto Health Care System who evaluates service members from all branches of the military. These interviews revealed the variations in the IDES documentation practices of different service branches. In early 2013, for example, the Navy used Disability Benefit Questionnaires (DBQs), which are problem-specific assessment instruments whose component questions are designed to elicit the information needed to complete a disability rating based on the rating schedules of Code of Federal Regulations Part 4. In early 2013, the workflow in Army's IDES did not use DBQs. However, according to members of our Advisory Board, all branches subsequently have transitioned to the use of DBQs. The current Separation Health Assessment (SHA) makes use of a General Medical (Gen Med) Examination DBQ template, which is intended to be a brief clinical summary and which requires that the details of each condition to be recorded in the individual specialty DBQs.

In response to these developments, we have focused our attention on DBQs as potential instruments for capturing structured functional assessment information. We developed semantic models of typical DBQs and investigated the nature of data elements in DBQs.

2. Propose a structure for CFA-CL. The CFA-CL coding scheme will be described in a document and also modeled as an OWL ontology using the Protégé tool.

In our original proposal, we hypothesized that CFA-CL codes will have the form of NNNN.e.xxxx, where *NNNN* is an ICF stem code at either the 3 or 4 digit level, *e* is an optional extension code that augments the ICF code to have greater specificity than that which is available in ICF, and *xxxx* denotes a set of category-specific qualifiers. For example, the Hip and Thigh Conditions DBQ evaluates the function of the hip in terms flexion, extension, abduction, and rotation. In ICF, the closest code for these functions is b7100 (functions of the range and ease of movement of one joint). We hypothesized that the stem code 7100 can be augmented by a 4-value extension code that indicates which of the more specific functions is being evaluated. For this extended stem code, we can use three qualifiers: the first indicating the specific joint that is involved (hip, in this case), the second indicating the laterality, and the third indicating severity, where the severity still depends on the specific function being evaluated.

Our detailed investigation of DBQ data elements suggests that developing a specific coding scheme from the outset is a suboptimal approach. First, a coding scheme is a syntactic construct and the optimal syntax is often dependent on specific use cases. For example, DBQs are organized in terms of data elements (e.g., "deep tendon reflexes of right knee") and data values (e.g., values from a 5-valued scale). From the point of view of capturing DBQ data, it is necessary to formulate the data as consisting of a data-element description and an acquired value, instead of coding it as a stem code with qualifiers. The key is that, if we have a consistent *semantic model* of the data elements and values, we can serialize them in alternative, equivalent syntaxes and infer the equivalence of data encoded in the different syntaxes.

Similarly, instead of creating arbitrary extensions to ICF codes, we can express the information content of the extensions more effectively as part of a semantic model of the data element. We need to distinguish between *measurements* of functions, where the entities being measured are often represented by external terminologies, and *assessments* that are abstractions conceptually closer to the notions that ICF codes are designed to represent. Logical Observation Identifier Names and Codes (LOINC), for example, has many ready-made codes for the measurements that are recorded in DBQs (such as extension, rotation, and flexion of various joints). Assessments, such as impairment of movement of the back (thoracolumbar spine), can be mapped to ICF. In both cases, detailed coding requires that we add additional qualifiers, such as whether a range of motion is measured after repetition of motions. Such qualifiers can be added as attributes in a semantic model of the data elements. Once we have a formal model of the data elements, if necessary, we can design a syntax for extension codes.

Given the difficulty of obtaining de-identified data for development purposes, and the impracticality of abstracting functional assessment information from directly narrative text, we consulted with some of our SAB members. We came to the conclusion that the best way forward is to focus, not on existing unstructured data and a syntactic coding scheme such as the one in the original project proposal, where functional assessment information is represented by a code stem and a set of code-specific qualifiers, but on developing a **framework for structuring and using functional assessment information** *prospectively*. This framework includes ontologies that describe the semantics of functional and related data elements, their relationships to standard terminologies and classifications, models of data-collection instruments, and data models for structuring assessed functional assessment information. We implemented the framework as a collection of ontologies using the Protégé tool. See Task 6 for details of how the semantic model for functional assessment information is structured.

The Protégé tool with which we create the ontologies and data models has a feature to export the content of the as a collection of inter-related HTML pages (Figure 3). This feature allows us to integrate documentation for the model as part of the ontology.

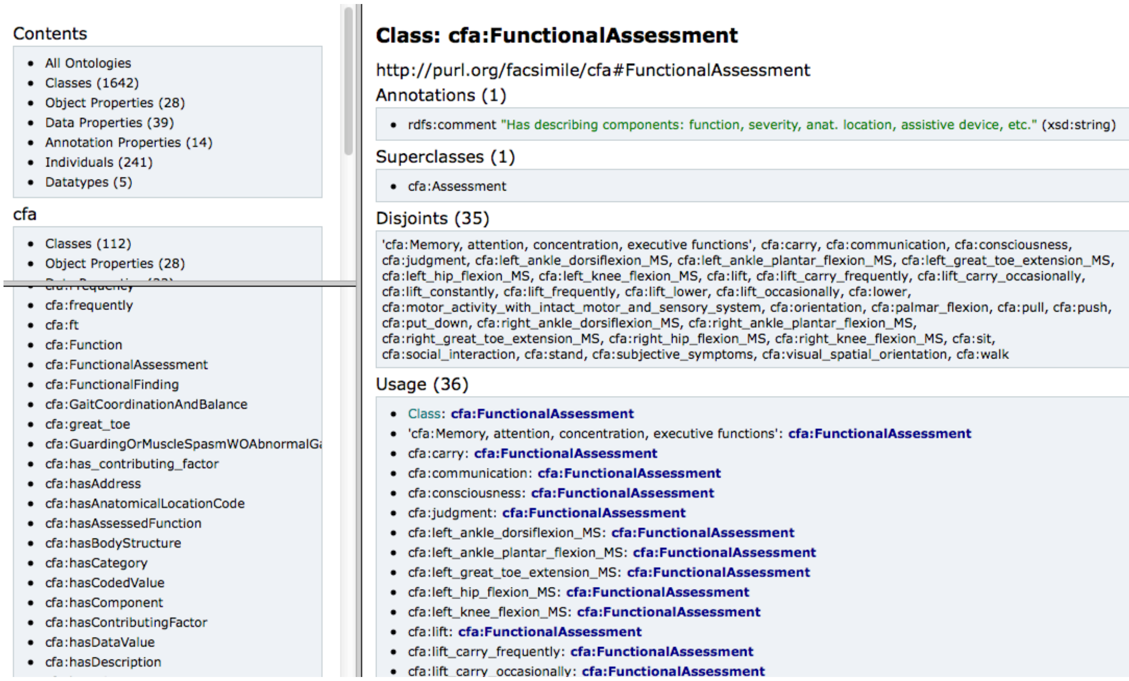


Figure 3. HTML pages documenting the CFA-CL ontology.

3. Using the proposed CFA-CL structure, we will develop a web-based editing tool for specifying CFA-CL code stems, their qualifiers, and value sets for the qualifiers.

We have successfully imported the CFA-CL semantic model into WebProtégé, which provides us with a Web-based environment for editing the ontologies, data models, and service-member data (Figure 4).

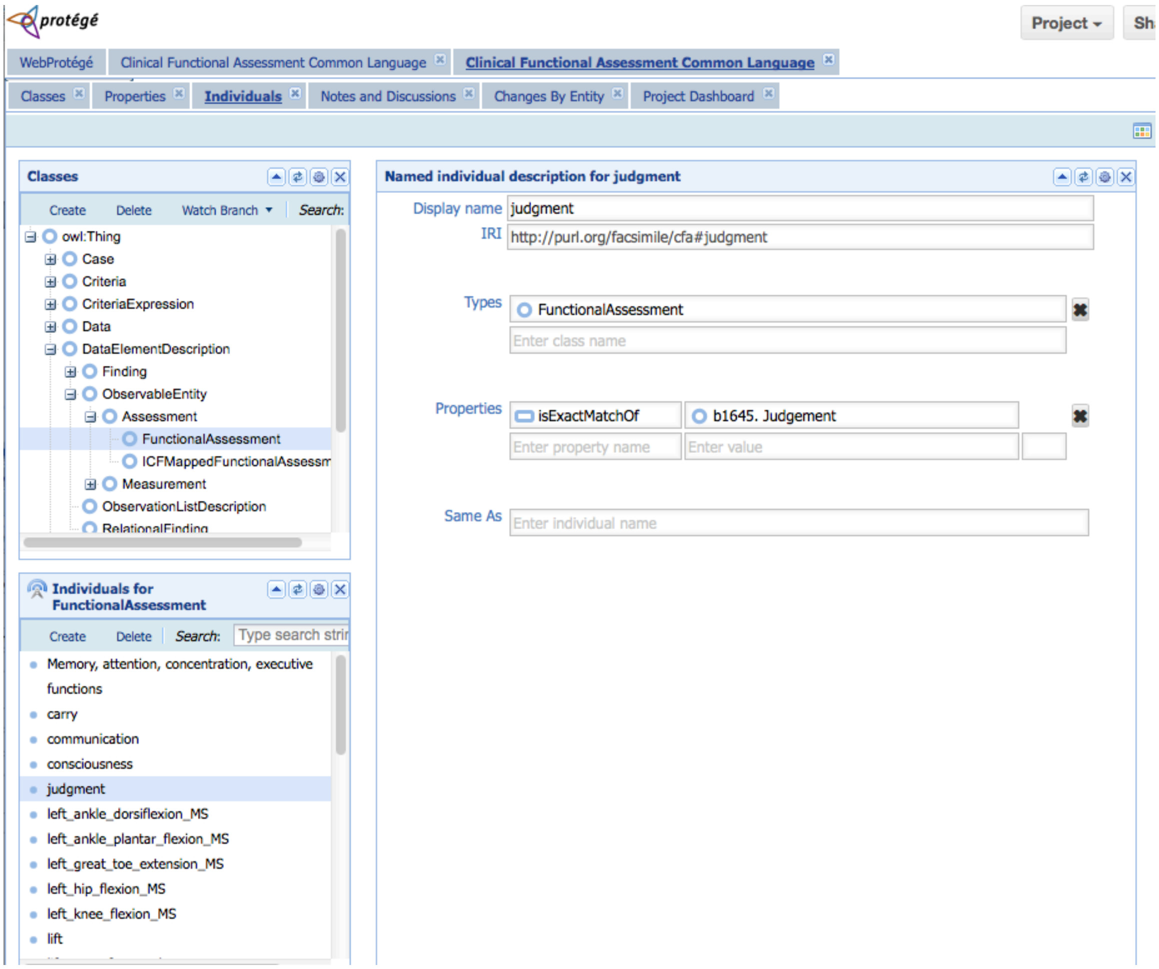


Figure 4. A WebProtege form for displaying and editing the "Judgment" data element. Note that this data element has an exact match to the ICF domain b1635.

4. We will populate CFA-CL with a selected subset of possible musculoskeletal code stems, their qualifiers, and value sets for the qualifiers.

We examined a set of existing instruments, including DBQs for lower back, knee and lower leg, ischemic diseases, and traumatic brain injury; the Military Occupation Specialties book; and SSA’s residual functional assessments.2 For the reasons discussed previously, we focused on the semantics of the associated functional entities and did not experiment with a specific coding syntax.

5. We will create a prototype CFA semantic model in which categories of impairment are defined by constraint expressions consisting of the CFA-CL and ICF code stems, qualifiers, and qualifier values.

In the CFA-CL framework, patient-specific functional-assessment data would be represented as semantic structures that are derived automatically as part of an enhanced data-entry process. We examined a set of existing instruments as described in Task 4 and modeled the structure and data elements in these instruments. Assessment instruments have sections and questions whose answers may be free text or may come from specific

2 The DBQs are VBA-21-0960M-14-ARE-Back.pdf, VBA-21-0960M-9-ARE-KneeLowerLeg.pdf, VBA-21-0960A-1-ARE-ischemic, NEURO - TBI Initial DBQ 9-15-11.doc.

value sets. Questions have descriptions of the data being solicited. Descriptions of questions and the value sets for their answers may use domain-specific terminologies.

To illustrate the structure of the CFA-CL Semantic Model, we take a question from the DBQ for the lower back. One of the assessments is a measurement of the forward flexion of the back (Figure 5):

SECTION IV - INITIAL RANGE OF MOTION (ROM) MEASUREMENTS

4. MEASURE ROM WITH A GONIOMETER, ROUNDING EACH MEASUREMENT TO THE NEAREST 5 DEGREES. DURING THE MEASUREMENTS, OBSERVE THE POINT AT WHICH PAINFUL MOTION BEGINS, EVIDENCED BY VISIBLE BEHAVIOR SUCH AS FACIAL EXPRESSION, WINCING, ETC. REPORT INITIAL MEASUREMENTS BELOW.

NOTE: Following the initial assessment of ROM, perform repetitive-use testing. For VA purposes, repetitive-use testing must be included in all exams. The VA has determined that 3 repetitions of ROM (at minimum) can serve as a representative test of the effect of repetitive use. After the initial measurement, reassess ROM after 3 repetitions. Report post-test measurements in section 5.

A. SELECT WHERE FORWARD FLEXION ENDS *(normal endpoint is 90)*:

☐ 0 ☐ 5 ☐ 10 ☐ 15 ☐ 20 ☐ 25 ☐ 30 ☐ 35 ☐ 40 ☐ 45
☐ 50 ☐ 55 ☐ 60 ☐ 65 ☐ 70 ☐ 75 ☐ 80 ☐ 85 ☐ 90 or greater

Figure 5. DBQ Range of Motion Measurement

We model the question as having a *focus* (i.e., the data element description), text, and possible values (Figure 6):

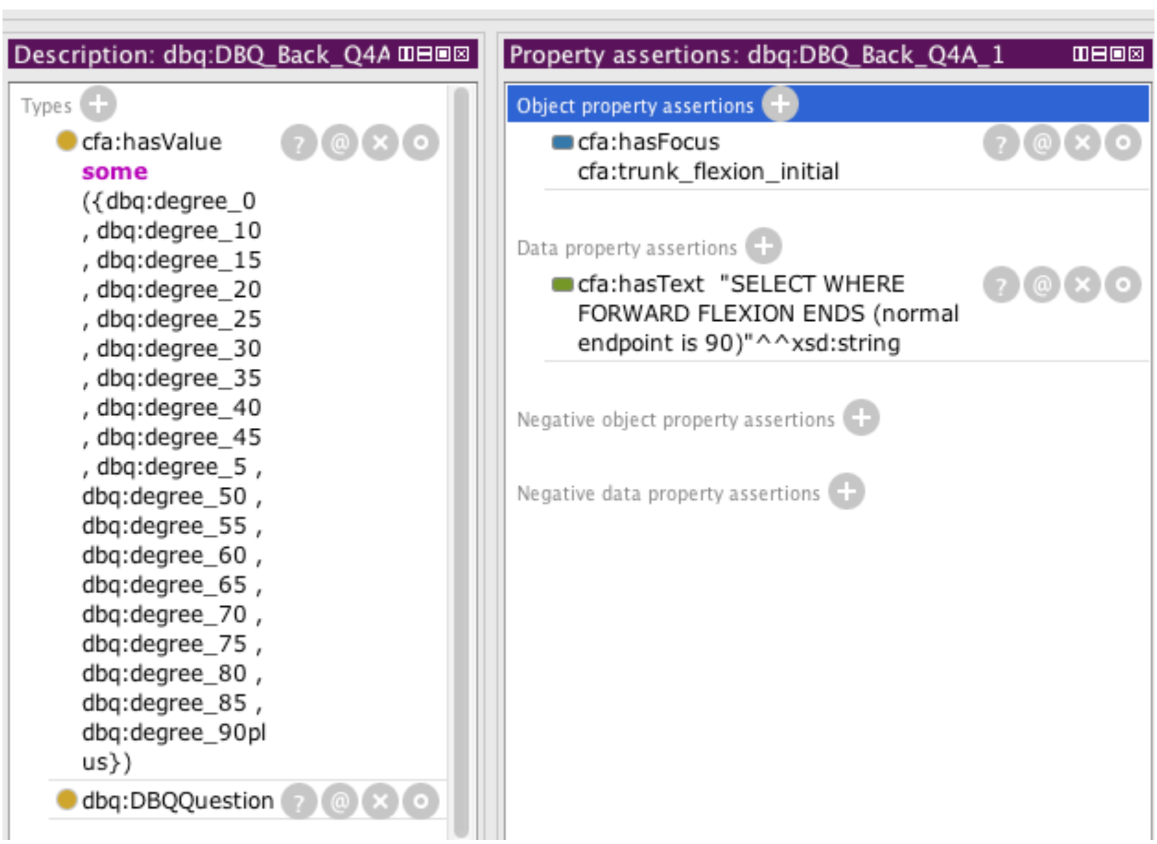


Figure 6. Modeling of Range of Motion data element.

The semantic model of the initial trunk-flexion data element is described in terms a number of properties (Figure 7):

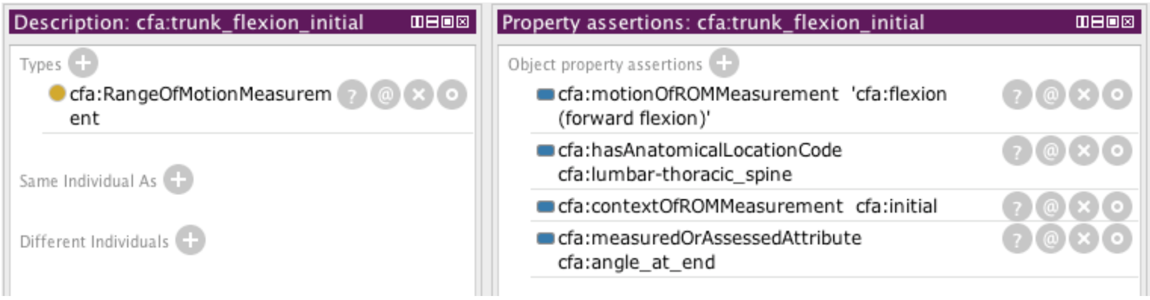


Figure 7. Semantic model for the trunk-flexion data element.

By associating the structured representation with components of an assessment instrument administered electronically, the acquired data can be converted as instances of CFA-CL Semantic Model automatically, obviating the need to have human reviewers extracting and coding the data. A structured datum representing an initial trunk flexion measurement would look like an EHR datum (e.g., an Observation in the Health Level 7 Reference Information Model). At the minimum, it would have a reference to the focus of observation (e.g., trunk flexion initial), a value (e.g., 80 degrees), and the ID of the patient.

Our analysis of the data elements in assessment instruments suggests that multiple terminologies are needed to formalize the data-element descriptions. DBQs explicitly require the use of ICD for coding diagnoses. Many signs and symptoms are concepts better coded in standard clinical terminologies such as SNOMED CT. Among functional assessments, a significant subset involves detailed measurements such as assessments of the range of motion in specific joints. ICF, with its relatively high-level functional categories, is not designed for recording such measurements. We have determined that, among standard clinical terminologies, LOINC has the appropriate codes for such measurements. For example, LOINC 41343-5 represents quantitative measurement of the angle of left-knee flexion. Currently, ICF is one of four standard terminologies to which we map descriptions of assessment-data elements. The mappings may be refined to specify that the data element description is an exact match, a specialization, or a generalization of the terminology concept.

6. We will define mappings between a selected subset of CFA-CL terms and ICF terms such that the mappings allow us to programmatically translate CFA-CL–coded data into corresponding ICF-coded data.

We model CFA-CL–coded data as instances of a class called datamodel:Observation (Figure 8). For ICF-coded data, we create a model consisting of classes that correspond to each of the four ICF axes: Body Structure, Body Function, Activities and Participation, and Environmental Factor. For each class, we specify the allowed qualifiers. Figure 9 shows the data model for “Body Structure” data. It specifies that an ICF-coded datum for body-structure impairment must include the nature, extent, and location of impairments.

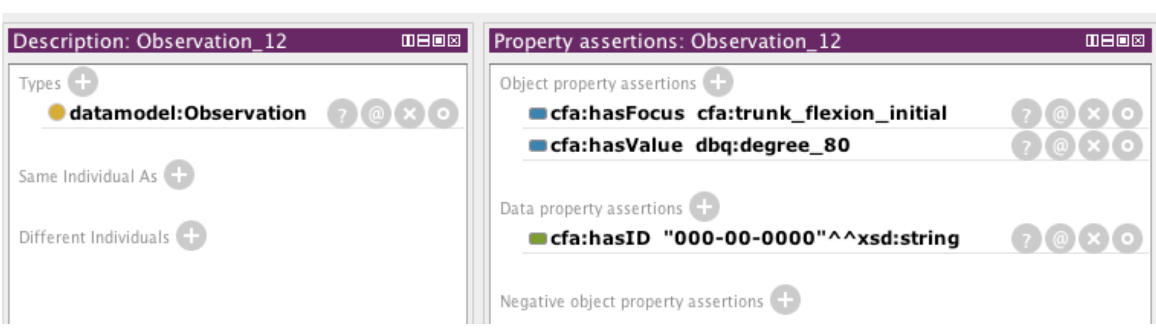


Figure 8. Modeling of collected data as an Observation.

Figure 9. Model for "Body Structure" ICF-coded data.

To facilitate the use of Semantic Web Rule Language (SWRL) rules to perform the mapping from CFA-CL to ICF, we first create a mapping structure CFA2ICFMapping (Figure 10), where a CFA entity (e.g., an assessment term or a qualifier term) is mapped to the corresponding ICF category or qualifier value.

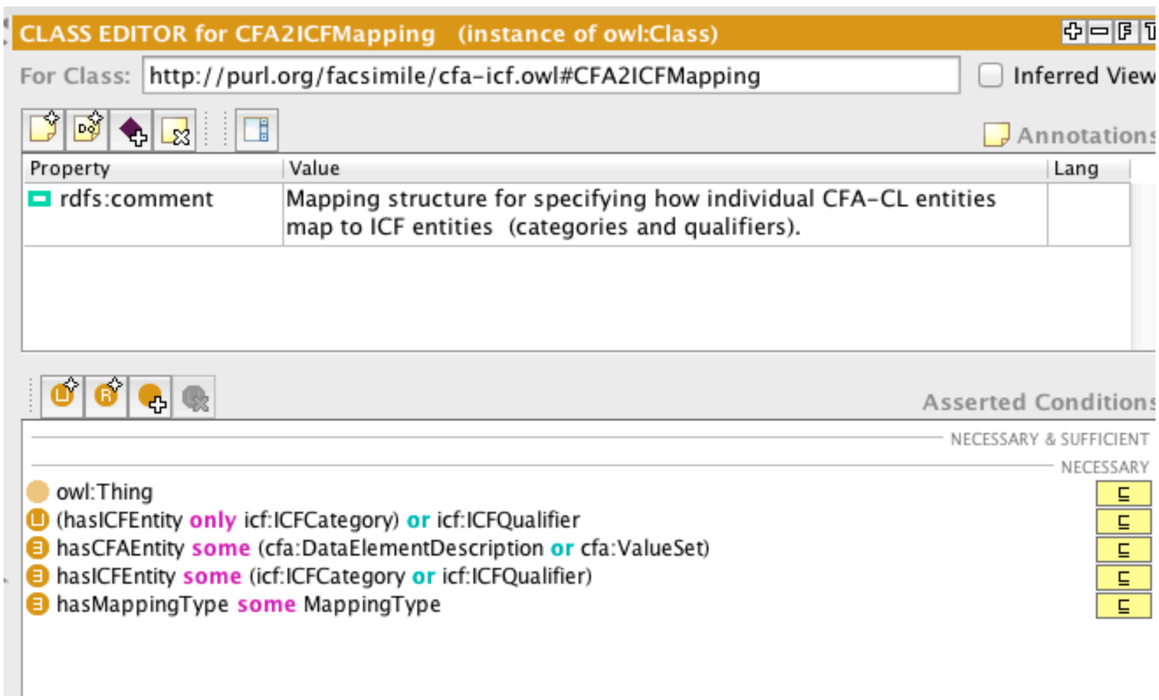


Figure 10. Mapping structure for translating CFA-CL coded data to ICF-coded data.

We model the mapping from the CFA-CL data format to the ICF data format as a collection of SWRL rules. An example is shown in Figure 11. It takes an Observation instance that encodes a body-function assessment (e.g., the right knee extension of 5 degrees) and mappings from CFA-CL to ICF (e.g., the notion of *extension* to “b7100 Mobility of a single joint” and that of 5 degrees to “3. SEVERE impairment (high, extreme, ...) 50–95 %” to create an instance of ICFBodyFunctionCode that denotes the mapped values as the combination of the ICF category and the ‘extent of impairment’ qualifier. To create the new ICF-coded data, we have to use Protégé’s SWRL extension built-in swrlx:makeOWLIndividual, which is not one of the standard SWRL built-ins.

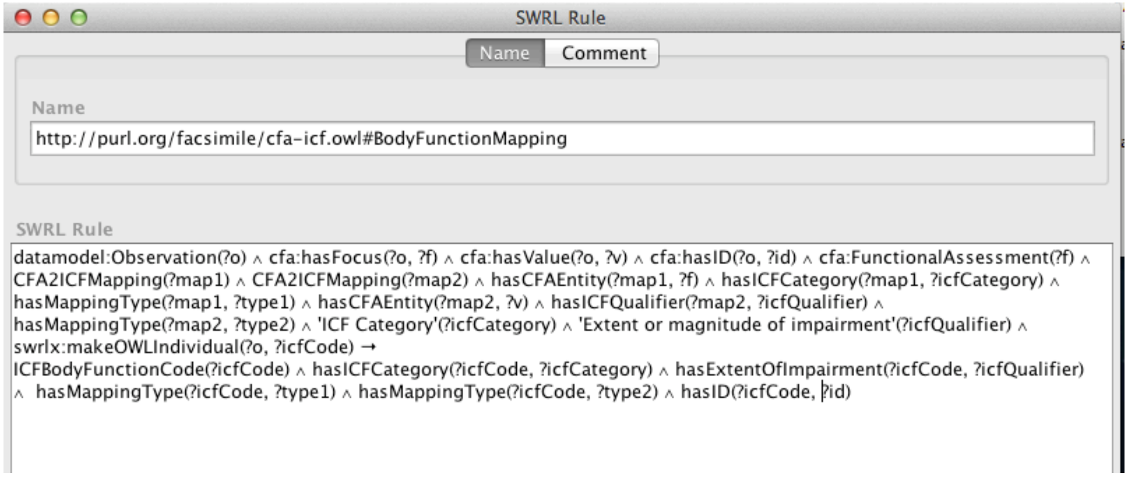


Figure 11. A SWRL rule for creating ICF-coded data from CFA-CL-coded body-function data.

Because ICF uses multiple codes to encode a single disability, we need to write additional rules to translate the CFA-CL data to a set of ICF codes. For the example of “right knee extension” observation, we use a second rule to generate the ICF body structure code (Figure 12). Note that we use an observation ID to indicate that the ICF body function and body structure codes are derived from the same CFA-CL observation.

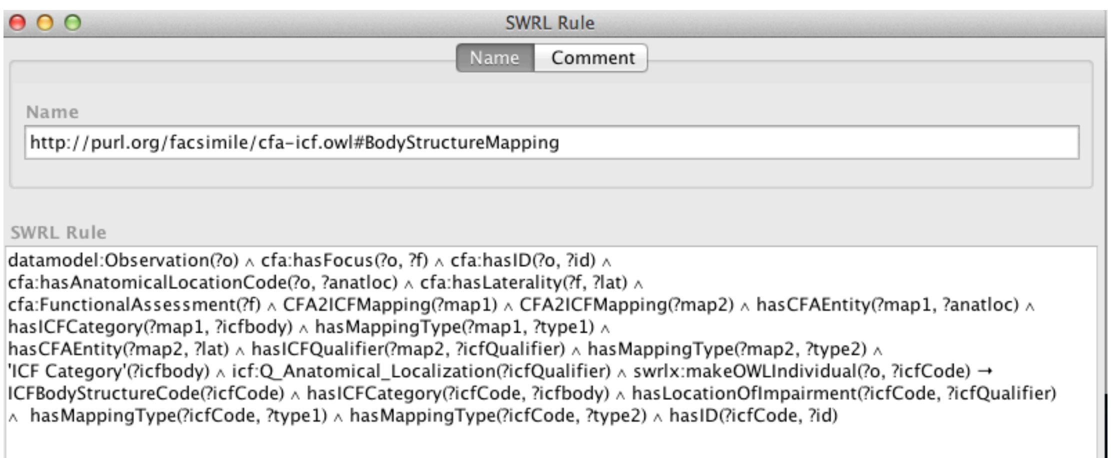


Figure 12. A SWRL rule to map the anatomical location of a body function assessment to ICF code.

7. We will create a developmental de-identified data set that contains musculoskeletal functional assessment. We will code the functional assessments using CFA-CL, and translate them to ICF-coded data.

As detailed in our report for Task 1, it is impossible to create de-identified data sets from the Madigan Army Medical Center archive. We are developing the CFA-CL for the possibility of capturing data prospectively and we are not relying on the availability of de-identified data retrospectively.

8. **We will define a set of queries that are interesting from the perspectives of evaluating individuals and of performing aggregated analysis. We will demonstrate the ability to make these queries on the developmental data set.**

Given our focus current focus on the DBQs, the queries that are most interesting from the perspective of evaluating individuals involve criteria from the Schedule for Rating Disabilities used in IDES to determine a numeric disability rating for the purpose of calculating the disability benefit. The criteria in the Rating Schedule are closely tied to questions in the DBQs. With our modeling of DBQ questions, we will be able to answer such queries.

From our interviews at Madigan AMC, we came to the conclusion that Madigan providers are intensely focused on the evaluation of individual service members, and have little interest in queries of aggregated data.

With data coded in the CFA-CL Semantic Model and mapped to ICF format, we can make aggregated queries such as *most common disabilities at three-digit level associated with amputation of the lower leg* or *disabilities associated with ICF coding s7501.418.* We can aggregate disabilities to any level and sort by frequency. With these queries, we may identify prevalence of specific problems (e.g., foot problems) that can be ameliorated with better equipment (e.g., change shoes, different inserts).

If we link CFA-CL data with other data sets, we can do much more interesting queries. For example, with appropriate data sets, we can mine for associations between functional assessments and the risk of homelessness after discharge or between amputation and incidence of diabetes. We can identify the need for home support (e.g., the need for aid and attendance) based on functional losses. These possibilities indicate the potential for using structured functional assessment data, but creating such data sets remains a daunting problem.

9. **We will specify the IDES task for which we will demonstrate the use of CFA-CL–coded data.**

We have analyzed criteria in the retention and rating standards. While many of these criteria, such as those related to range of motion, can be matched precisely from structured data, others, such as "Loss of toes that precludes the abilities to run or walk without a perceptible limp and to engage in fairly strenuous jobs" require subjective judgment to identify. We may not be able to match such criteria with the data collected through any assessment instruments. Therefore, we believe that at most we can index the criteria with relevant codes, and that we can use the coded data for an individual subject, once the data become available, to focus attention on those criteria that may be relevant to that subject.

Key Research Accomplishments

We have come to the conclusion that current documentation practices in centers such as Madigan AMC present difficulties in codifying clinical functional assessments as structured data.

We have identified as a problem a lack of structured functional assessment data because there is no standard data representation that is in use. Yet the representation we are creating is difficult to evaluate because of the lack of data. A strategy to break the chicken-and-egg problem of data representation and data capture is to instrument systems for entering form-based data so that, as data are entered into forms, they are automatically transformed into our underlying models.

We found that DBQs, the criteria in the MOS Manual and in the military retention standards require very specific functional assessments, which are difficult to map to ICF. There is no Rosetta Stone for translating neatly among these different data elements. What we can accomplish is to create a set of models that provides a mechanism for representing diverse data related to functional assessment.

We found that, when the goal is to automate the capture of structured functional assessment data, the particular syntax that we initially proposed is not necessary. Instead the data can be structured in a semantically sound representation that facilitates queries and transformations.

For the goal of capturing structured clinical functional assessment, we have created a semantic model of data element descriptions and a framework for using these descriptions to structure data.

Conclusion

We have created a semantic framework for modeling structured functional-assessment data and showed how such data can be derived from assessment instruments such as DBQs. We have created mapping structures and rules to transform data represented in this framework to ICF-coded format.

We have not been able to determine how such structured data can be derived from an existing IDES workflow that relies primarily on PDF documents that cannot easily be parsed or de-identified.

PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS

Nothing to report

INVENTIONS, PATENTS AND LICENSES

Nothing to report

REPORTABLE OUTCOMES

We have created a semantic model for clinical functional assessment consisting of

- An ontology of functional assessment data element descriptions
- An information model of assessment instruments and its components
- A data model for assessment data, in CFA-CL and ICF formats

We showed how data represented in this form can be mapped into ICF-coded format.

OTHER ACHIEVEMENTS

Nothing to report

References

None.

Appendices

Functional Assessment Coding Using International Classification of Functioning, Disability, and Health (ICF)

ICF is a multipurpose classification that, together with International Classification of Diseases (ICD), is a reference classification in the WHO Family of International Classifications (WHO-FIC). It provides a standard language and conceptual basis for the definition and measurement of functions and disability. Unlike a medical model of disability, which sees loss of functions only as consequences of diseases and disorders, ICF embodies a "bio-psycho-social synthesis" that conceptualizes function and disability in the context of health conditions, environmental factors, and personal factors (Figure 13).

We first introduce ICF's component structures, then we examine ICF's coding scheme to evaluate it as a candidate common language for coding functional-status information required in the disability-evaluation process.

1.1.1 ICF Structural Components

Structurally, ICF organizes information in two parts: (1) *Functioning and Disability*, which comprise the body functions (*b* codes), body structures (*s* codes), and activities and participations (*d* codes), and (2) *Contextual Factors*, which include environmental factors (*e* codes) and personal factors, which have not been developed systematically yet. Each component consists of various *domains* (e.g., *d4 Mobility*); each domain, in turn, consists of *categories* (e.g., *d450 Walking* and *d4500 Walking Short Distances*), which are the units of classification. Figure 13 depicts the interactions among the components of ICF.

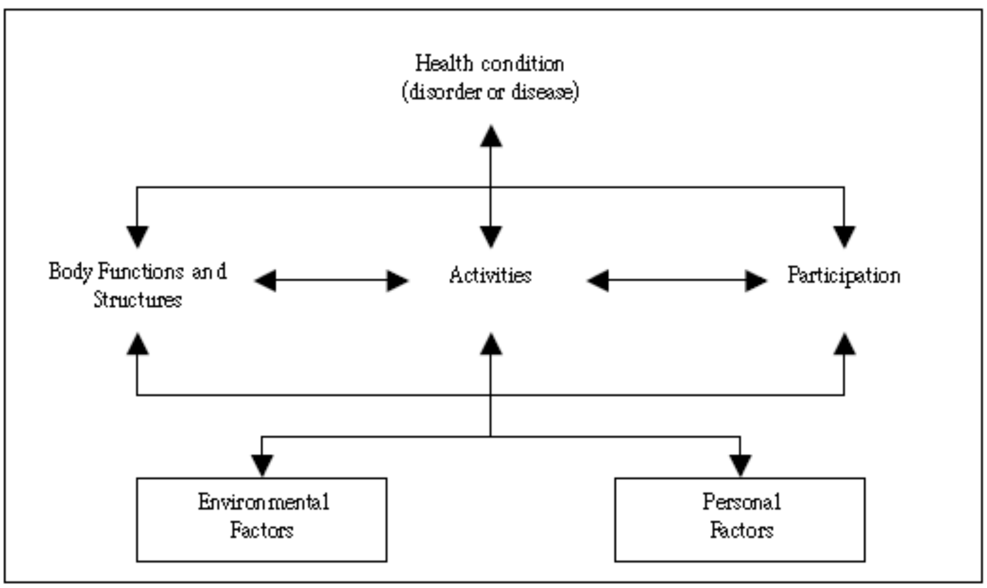


Figure 13. Individual's functioning in a specific domain is an interaction or complex relationship between the individual's health condition and contextual factors (i.e., environmental and personal factors). In one direction, health conditions have impact on body functions, body structures, and the capacity to perform activities or to participate in economic, social, and civic life. Personal and environmental factors facilitate or restrict functions and capacities. Conversely, the presence of disability may modify the health condition and a person's activities may modify environmental factors (Source: [4] p. 18).

1.1.2 ICF Coding

Coding different health and health-related states requires that ICF categories are used in conjunction with component-specific *qualifiers*. The body structure component, for example, requires three qualifiers that specify the extent of impairment (first qualifier), the nature of impairment (second qualifier), and location of impairment (third qualifier). Each qualifier has a generic value set. For example, the extent-of-impairment qualifier for the body structure component has the following value set:

Table 1. Generic severity scale

0	NO impairment
1	MILD impairment
2	MODERATE impairment
3	SEVERE impairment
4	COMPLETE impairment
8	not specified
9	not applicable

In another example, the activities-and-participation component has two default and two optional qualifiers as depicted in Figure 14.

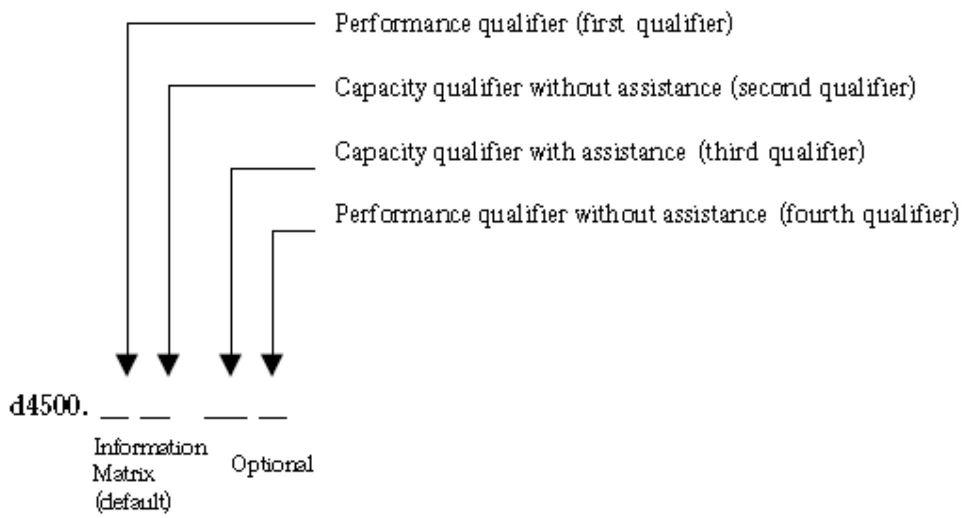


Figure 14. The default and optional qualifiers of the activities and participation component. Source: [4] p. 230.

The ICF reference document ([4] p. 123) says that:

> The performance qualifier describes what an individual does in his or her current environment. ... The capacity qualifier describes an individual's ability to execute a task or an action. This construct aims to indicate the highest probable level of functioning that a person may reach in a given domain at a given moment.

The use of ICF components and component-specific qualifiers gives the ICF coding scheme tremendous *post-coordination* capability, where complex situations are described by selecting appropriate ICF categories and by qualifying them with additional codes. If ICF were to pre-coordinate, that is, pre-enumerate and define, all possible combinations of ICF categories and qualifiers, the classification would explode into a gigantic and unusable tree. ICF encourages the use of as many coding constructs as necessary to represent a health or health-related situation. For example, *Able to walk for more than 1 mile with right leg impaired and with hand-held assistive device (HHAD) in one hand*, can be coded with the following combination of codes:

d4501.880: Walking long distances (more than a kilometer) (*d4501*), unspecified performance qualifier (*8*), unspecified capacity without assistance qualifier (*8*), no impairment in capacity with assistance (*0*)

s7501.881: Structure of lower leg (*s7501*), extent and nature of impairment not specified (*88*), location of impairment: right side (*1*)

e1151. +8: Assistive products and technology for personal use in daily living (*e1151*), facilitator not specified (+8).

This requirement to use multiple codes (and their qualifiers) to code different dimensions of a functional assessment means that it is possible to aggregate data along these dimensions. For example, disability of the lower extremity may have multiple functional consequences.

Coding Experiment with SSA Disability Evaluation

In 2011, the Stanford team completed IPA contracts with the Social Security Administration (SSA), through which we investigated the coding requirements for SSA disability determination and methods for mapping a prototype SSA coding scheme to ICF. Our experience suggests that it is possible to develop a CFA-CL that uses ICF categories (e.g., “d4501” — walking long distance) as code stems to which coders append category-specific qualifiers (e.g., distance walked, laterality of impaired lower leg) to represent the severity and location of the impairments. We model the value sets for the category-specific qualifiers as constraints on relevant ICF coding. (Thus, it would be impossible to indicate “laterality of impaired lower limb” in the context of a category such as “Shortness of Breath.”)

The primary goal of the SSA study was to create a coding scheme that is easy to use, that captures fully the functioning information described in the SSA disability assessment, and that maps rigorously to the ICF coding scheme. Fortunately, the mapping has to be done only once, and, once it is completed, allows automated translation from SSA data to ICF-compliant coding.

This approach includes the following components:

1. An SSA coding scheme that has (a) 3-digit stem codes that are analogous to the 3-digit ICF activity and participation codes, (b) 3-digit qualifiers that represent the capacity, localization, and environmental factors associated with the 3-digit stem codes.
2. Mappings to the ICF coding scheme where (a) the 3-digit SSA stem codes are mapped to combinations of ICF activity and participation codes and (b) the 3-digit SSA qualifiers are mapped to ICF capacity, severity, body structure, and environmental-factor qualifiers.

We illustrate this coding approach with the following examples:

Table 2. Sample Hypothetical SSA Coding

	Functional Assessment	Hypothetical SSA Coding
1	Able to walk for ~ 0.4 km with right leg impaired and with HHAD in one hand	450.211 (450:Ambulating; xxx.2: ~0.4 km; xxx.x1: right; xxx.x1: HHAD in one hand)
2	Able to lift and carry 20 lb occasionally or 10 lb frequently with right hand/arm impaired using HHAD in both hands	430.210 (430: lifting, carrying with upper extremity; xxx.2:20lb/10lb occasionally/frequently; xxx.x1:right;xxx.xx0: HHAD in both hands)
3	Able to grasp small objects, but limited fine control, with right hand impaired, using HHAD in one hand	440.211 (440:Fine movements of the upper extremity; xxx.2 Able to grasp small objects, but limited fine control; xxx.x1: right; xxx.xx1: HHAD in one hand)
4	Able to push and pull 20 lb occasionally or 10 lb frequently with right hand/arm impaired using HHAD in both hands	445.210 (445: Pushing or pulling with upper extremity; xxx.2: 20 lb occasionally or 10 lb frequently); xxx.x1:right; xxx.xx0: HHAD in both hands)

In this proposed SSA coding scheme, the meaning of concepts is tailored to SSA "Blue Book" listings [5]and Residual Function Capacity assessments. For example, the 3-digit code 445 (Pushing or pulling with upper extremity) has no exact equivalent at the 3-digit level in ICF. Other codes may have only an inexact match. For example, the SSA code 430 “Lifting/carrying with upper extremity” is not identical to ICF d430 “Lifting and carrying objects," because the latter includes the possibility of lifting and carrying objects on the head (d4304).

The 3-digit SSA stem code and the 3-digit Capacity, Localization, and Environment qualifiers have the form NNN.CLE. This compact notation captures very complex and specific information. The trade-off that we make here is the need for category-specific qualifier codes. Unlike ICF, where qualifiers are component-specific but otherwise generic, qualifiers in the proposed SSA coding scheme are code-specific. For example, SSA code 450 has distance-related capacity qualifiers, whereas SSA code 430 has weight-related capacity qualifiers. Furthermore, when used with code 450, the localization qualifiers refer to impairment of legs, whereas in the case of code 430, the localization qualifiers refer to impairment of upper extremities. Instead of a 15-page document, like the one developed by WHO to describe the coding guidelines for ICF, the SSA coding scheme requires a detailed description for each 3-digit category. Such a detailed manual that describes category-specific qualifiers is not particularly onerous to produce because the scope of SSA’s functional coding is much more limited than that of ICF.